## *Issues and Controversies in DILI*

| Session I: FDA Concept Paper – Premarketing Evaluation of DILI | | | |
|---|---|---|---|
| **Function** | | **Name** | **Affiliation** |
| Moderator | | Paul B. Watkins, MD | UNC |
| Recorders | 1.1 | Christine Hunt, MD | GSK |
| | | Joyce Korvick, MD | FDA |
| | 1.2 | Judy Racoosin, MD | FDA |
| | | Alan Goldhammer, PhD | PhRMA |

| Session 1.1 | |
|---|---|
| **Drafting of Concept Paper: Premarketing Evaluation of DILI** | **Ruyi He, MD (FDA)** |

Dr. He presented an overview of the concept document. He noted that when analyzing a liver safety database from clinical trials, a high level of liver safety concern results from finding any of the following three features:

1. a higher rate of aminotransferase (AT) >3xULN in drug treated patients than controls
2. marked elevations of AT (10, 20xULN), and especially
3. Hy's Law cases – drug treated patients with evidence of hepatocellular injury (AT>3xULN) and total bilirubin $\geq$2xULN and no other explanation for liver disease/drug.

Dr. He noted that the specificity of Hy's Law in predicting potential for severe liver injury appears to be high; there are no false positive cases known to FDA.

Dr. He then reviewed various issues outlined in the concept document. He ended by commenting that the FDA encouraged collection of blood and urine samples for genomic, proteomic, metabolomic assessments.

| **Rationale for Concept Paper and Proposed Guidance** | **John Senior MD (FDA)** |
|---|---|

Dr. Senior reviewed several issues pertinent to the concept document, including the fact that DILI is now the major cause of acute liver failure in the US. He also noted that Bob Temple coined the term "Hy's Law" based on the observation by Dr. Hyman Zimmerman that hepatocelllular jaundice due to a drug has at least a 10% mortality.

He then reviewed one way to graphically display peak liver chemistry values obtained in a clinical trial. The log peak serum ALT is plotted on the horizontal axis and the log peak serum total bilirubin is plotted on the vertical axis. The display can then be divided into 4 quadrants by the lines through ALT 3 X ULN and bilirubin 2 X ULN. The left lower quadrant is then considered normal or insignificant elevations in liver chemistries, the left upper quadrant indicates patients with cholestasis; the right upper quadrant are the Hy's Law patients; the right lower quadrant was termed Temple's Corollary (patients with ALT > 3 x ULN but not satisfying Hy's Law.

Dr. Senior then reviewed the importance of symptoms when evaluating liver signals. He noted that in a large study of 11,141 patients treated with isoniazid, there were no deaths and only 11

discontinuations when treatment was discontinued based on symptoms compatible with hepatitis (Nolan, et. al., JAMA 1999; 281: 1014-8). This is in contrast to a much higher rate of discontinuations and deaths in earlier studies.

Dr. Senior then reviewed a 1975 study where patients were continued on INH therapy despite high elevations of liver chemistries, including three patients who would be considered Hy's Law cases. In these patients, progressive liver injury was not observed despite the continued administration of INH suggesting that adaptation to continued drug exposure can occur even with considerable liver injury with impaired function.

Dr. Senior ended by reviewing potential areas for discussion prompted by the concept document, including the need for better causality assessment tools.

| **Industry Commentary on Concept Paper** | **Jerry Stern MD (Boerhinger Ingleheim)** |
|---|---|

The following suggestions were made:

1. It was noted that the target audience for the document is not hepatologists and will include non-clinicians. The document should therefore be more straightforward and easy to follow

2. The document would benefit from more background information relevant to drug induced liver injury, including discussion of routine liver chemistries and their interpretation, and relevant disorders such as Gilbert's Syndrome.

3. In the definition of a Hy's law case it should be noted that the bilirubin should be predominantly conjugated bilirubin. It was also suggested that "substantial increase" in alkaline phosphatase be specifically defined as >2xULN.

4. In Section 3.1, the recommendations should be modified for special patient populations, such as patients with diabetes, obesity/NASH, HIV, and chronic liver disease.

5. It was noted that recommendations should be different for early (phase 1-2) vs. late (phase 3) studies. More caution is appropriate when clinical experience with the compound is limited.

6. Only patients with well-characterized and stable liver disease should be initially included in clinical trials as unexplained elevations in baseline liver chemistries represent an undefined risk. In some populations, liver disease probably increases the risk of DILI

7. If symptoms suggest clinical hepatitis – study drug should be stopped even if defined liver chemistry thresholds are not exceeded.

8. If study drug is being administered for a non-life threatening indication, the defined criteria for stopping drug should be more conservative than for life threatening indications.

9. It should be clarified how long drug treatment should be continued with persistent ALT elevations 3-5xULN

10. Stopping rules should be stated for subjects with abnormal baseline liver chemistries

11. Sample CRF would be valuable to ensure consistent reporting of information for future FDA analyses

12. Free text boxes should be minimized as they allow non-standard nomenclature and are not amenable to programmable analyses (particularly in large global studies)

13. The lack of standardized DILI nomenclature complicates interpretation of signals and this should be addressed.

| **Industry Commentary on Concept Paper** | **Deborah Kirby, MD (Pfizer)** |
|---|---|

The following suggestions were made:

1. The document should acknowledge greater flexibility in designing studies based on known preclinical or clinical data, or known drug class effects.

2. The document's title should reflect the fact that scope is primarily "clinical"

3. The document would benefit from further discussion of adaptive vs. non-adaptive liver chemistries.

4. It will be important to continually update the final guidance document as knowledge grows

5. It will be important to explore ways to organize/leverage FDA information across drugs and companies to maximize learning.

6. The term "more frequent" should be defined in reference to comparison between treatment and control in the incidence of ALT>3xULN.

7. It should be stated that ALT is the primary parameter to assess hepatocellular injury.

8. A discussion of the use of the ALT/AST (alcohol) would be helpful.

9. INR/PT is not consistently mentioned in the document in terms of assessing liver function. Perhaps the following statement should be made: "obtaining additional tests (e.g. PT/INR), may be appropriate."

10. More detail should be given to the frequency of retesting in the face of elevated liver chemistries.

11. Separate criteria should be provided for subjects with elevated baseline ALT

12. Recommendations concerning the duration of follow up at the end of the study should be stated – 4 weeks might be reasonable.

13. Standardized case report forms should be encouraged.

| AASLD Commentary | Neil Kaplowitz MD (USC) |
|---|---|

Dr. Kaplowitz raised the following issues and questions related to the concept document:

1. Once a case of allergic DILI is observed [e.g. propylthiouracil (PTU), Augmentin, Aldomet, nitrofurantoin, trovafloxacin], should the criteria for monitoring and stopping treatment be modified?

2. The document largely ignores cholestatic liver disease, but recent reports suggest significant mortality in cholestatic injury (although less than observed with hepatocellular injury).

3. There is little data to guide how mixed type liver injury should be treated.

4. Probably important to exclude patients with advanced liver disease in early clinical trials, but fairly advanced cirrhotics can have normal liver chemistries. Abdominal ultrasound may be required to detect these people.

5. Perhaps the document should give some guidance on when premarketing data would suggest the need for post marketing monitoring of liver chemistries.

6. Muscle injury can result in ALT elevations so CPK measurements should be suggested.

7. The definition of rapid ALT elevations in the document should be more carefully defined.

8. It is clear that much is unknown and research on DILI should be stimulated. Particular areas of emphasis include:
    - Genetics – drug metabolism, innate and adaptive immunity
    - Biomarkers – we should develop more; early predictors and specificity – especially time course
    - Serum adducts to assist causality assessment (e.g., acetaminophen hepatotoxicity)
    - Role of inhibition of the bile salt export pump (BSEP) – should serum bile acids be monitored in clinical trials if BSEP inhibition is observed *in vitro*?

| NIH/DILIN Commentary | Paul Watkins, MD (UNC) |
|---|---|

Dr. Watkins stated that the document had been circulated to the principal investigators of the DILIN network (Drs. Herbert Bonkovsky, Naga Chalasani, Timothy Davern, Paul Watkins, and Robert Fontana), and NIH staff involved with the network (Jay Hoofnagle, Leonard Seeff and Jose Serrano). A conference call was then held and consensus was achieved on the following points:

1. The document creates a new term “Hy’s Law signal” to describe the presence in a liver safety database of both Hy’s Law case(s) and excess ALT elevations > 3 X ULN in

treated vs. controls. This is clearly the usual setting for Hy's Law case, but readers may not understand that the presence of Hy's Law cases is a worrisome signal even in the absence of background ALT elevations (e.g., in early clinical trials where the number of treated patients may not be sufficient to detect a true difference between control and treatment in incidence of ALT elevations). There may also be confusion between the definitions of Hy's Law Signal and Hy's Law. The DILIN investigators recommended dropping the term Hy's Law Signal.

2. 2). In phase three studies, the DILIN investigators recommended raising the ALT level for stopping treatment in patients with asymptomatic ALT elevations to 10 X ULN ( from 8 X ULN). This assumes very close monitoring of all patients with ALT > 5 X ULN.

3. 3). Enrollment of subjects with abnormal ALTs should be encouraged in clinical trials, but the pretreatment ALT should generally not exceed 3 X ULN. Treatment algorithms should then generally employ fold-baseline in place of fold ULN. However, the discontinuation level of asymptomatic ALT elevations in phase 3 trials should remain 10 X ULN.

4. 4). Non-alcoholic steatohepatitis (NASH) should be mentioned in the document as a cause for background liver disease and as a condition that can be exacerbated or perhaps caused by some drugs (e.g. tamoxifen).

5. 5). Set liver chemistry values should be considered as cut offs rather than multiples of the ULN. Networks such as DILIN have found little variation in the actually values obtained for serum transaminase measurements across different labs, although ranges of normal vary up to 2 fold among the same labs. At the very least, labs used in clinical trials should provide the performance characteristics on widely used standardized samples.

| **Open Discussion prior to AM Break (Session 1.1)** | **Various Identified Below** |
|---|---|

**Jim Freston MD (U. Conn)**
The document does not discuss causality assessment processes. For example, it might be helpful to provide guidance on when companies should consider having a panel of hepatologists assess a given case. Furthermore, if the panel does not consider a given case to be drug related, is the case no longer "counted" in assessing liver safety? An opinion from the FDA was requested.

**John Senior MD (FDA/CDER)**
There is no standard process or rule to adjudicate cases. The concept is to get consistent data from all cases, so that they can be fully evaluated.

**Bob Temple MD (FDA/CDER)**
One problem with expert opinions is that they can change over time. For example, with ticrynafen there were 7 cases of liver injury in the NDA database and none were assessed by experts as related to the drug. For example, it was noted that one case was from an area where there had been a "hepatitis outbreak"; others were attributed to various other causes. However, there were no cases in the placebo-treated group and, in retrospect; the majority of

cases were clearly due to the drug. The control group is therefore very important. Events occurring in the post-marketing period are obviously much harder to assess.

**Ross Pierce MD (FDA/CBER)**
Expert adjudication may not give always give an accurate answer. In a large statin dose-response study, we asked the sponsor to do a dose-response analysis of the incidence of aminotransferase elevations, *limiting the analysis only to those subjects whose enzyme elevations had been adjudication (possibly blinded) as not related to statin administration*. The analysis indicated a positive dose-response trend with a p value of 0.6, suggesting that many of the expert adjudications had been inaccurate in their conclusion of a lack of causal relationship between the test statin and the aminotransferase elevations.

**Mark Avigan MD (FDA/CDER)**
Causality assessment can be very difficult in patients with underlying liver disease. For example, immune reconstitution during treatment of HIV infection can result in a flare of liver injury in patients chronically infected with viral hepatitis. This appears clinically as DILI but actually is not.

**Bob Temple MD (FDA/CDER)**
The background incidence of ALT > 3xULN is also helpful in evaluating a single case of severe liver injury. If the incidence of ALT elevations > 3xULN is not different in treated and control subjects, it is more likely that a single severe liver injury is unrelated to drug. If such a case occurs in the setting of an increased incidence of ALT >3xULN in the treated group, the case is more likely due to drug.

The concept underlying Hy's Law is that the increase in bilirubin is the result of a loss of functioning hepatocytes and not cholestasis. If the AP is elevated substantially, it is not a pure Hy's Law case.

**Paul Watkins MD (UNC)**
If the bilirubin rise is secondary to loss of hepatocytes, then the ALT rise should precede the bilirubin rise. In an acute injury, the ALT should be in the thousands, rather than hundreds. Perhaps the 3xULN cut off for Hy's Law is too low.

**Bob Temple MD (FDA/CDER)**
This has been generally true. Drugs that have produced Hy's law cases have been associated with high ALT elevations (>10xULN).

**Neil Kaplowitz MD (USC)**
Does this include drugs with allergic injury?

**Bob Temple MD (FDA/CDER)**
We cannot look at preclinical databases to assess this.

**Neil Kaplowitz MD (USC)**
I'm not sure we should be calling something a Hy's Law case, when the bilirubin is only moderately elevated (and AT > 3xULN). I'd like to see higher bilirubin elevations in Hy's law cases. Also, I think we rely too much on AP – which doesn't always behave as you'd like it to.

**Bob Temple MD (FDA/CDER)**

AT is typically well above 3xULN in Hy's Law cases. The cut off of 3xULN is a practical one. If there is no ductal obstruction, elevation of bilirubin to 2xULN means that there has been sufficient liver tissue damage to yield clinical concern.

**Paul Watkins MD (UNC)**
It is said that there was no difference in the incidence of ALT elevations between treated and control patients in the felbamate clinical trials, and yet this drug has potential for severe liver injury. Are there other examples?

**Mark Avigan MD (FDA/CDER)**
Some drugs have multiple different signatures of injury and not all drugs associated with DILI cause high elevations of transaminases when liver injury occurs – e.g. Augmentin. FIAU produced fatal liver injury without ever causing a striking ALT elevation.

**Jack Uetrecht PhD (U. Toronto)**
The audience for this document is not hepatologists. Real patients are not simple – so when people use this document on real patients – they will get frustrated and tend to use the worst case scenario.

The document should give a sense that the reality is a lot more complicated.

**Paula Lapinskas (Celgene Corp)**
Clinical data suggests GGT and other markers may be useful. Any comments on other markers of hepatotoxicity?

**Neil Kaplowitz MD (USC)**
Some people have looked at alpha-GST as a liver marker, but it does not seem to have practical advantages.

**Paul Watkins MD (UNC)**
Does anyone feel there's a useful marker that should be discussed in the document?

**Herb Bonkovsky MD (U. Conn)**
GGT has high sensitivity and poor specificity. Bile acids are sensitive, but not specific. If bile salt secretion is affected, you will see bile acid elevation.

I believe that the concept paper would benefit from "bullet points" and providing a causation assessment algorithm to improve readability.

I also believe that exclusion of biliary tract disease should receive greater emphasis in the document.

**Jim Freston MD (U. Conn)**
A number of companies have developed excellent causation assessment algorithms.

**Jim Lewis MD (Georgetown)**
A lot of companies are gun-shy about continuing subjects with ALT>3xULN (due to concerns about interpretation of Hy's Law).

It is probably best to study patients with abnormal baseline liver chemistries in a separate study rather than simply adding a small number to larger studies.

**Leonard Seeff MD (NIH/NIDDK)**
ALT 8-10xULN is suggested as a threshold by the NIH DILIN group due to adaptation. It's not possible currently to assess whether adaptation will occur or injury will progress without continuing to treat patients through lower level elevations.

**Christine Hunt MD (GSK)**
Following on the point raised by Jim Lewis, is there any interest in gathering and sharing liver chemistry information for patient populations with elevated baseline liver chemistries? For example, the proposed liver chemistry subject stopping criteria would result in 14% or more of subjects with chronic hepatitis B stopping study drug, which might prove efficacious.

Additionally, some global ethics committees currently object to the use of ALT 5xULN as a liver chemistry subject stopping threshold. In studies of healthy volunteers, it will be challenging to get approval to treat through to elevations of ALT 8xULN or 10xULN.

**Becky Taub MD (Roche)**
This document really ties back to the need to understand the science underlying the adaptation phenomenon.

**John Senior MD (FDA/CDER)**
Another issue we do not understand is why there is a variable latency prior to onset of liver injury. It is unlikely that the mechanisms that underlie injury are the same for every drug.

**Bob Temple MD (FDA/CDER)**
A good example of latency is troglitazone where some cases occurred after many months and possibly more than 1 year of therapy. There has been a lot of research on the mechanisms underlying troglitazone DILI but no explanation for this latency has been provided.

| **Session 1.2** | |
|---|---|
| **Is the susceptibility to DILI in the genome?** | **Larry Lesko PhD (FDA)** |
| **Potential mechanisms of ximelagatran hepatotoxicity** | **T. Andersson PhD (AstraZeneca)** |

No summary recorded.

| | |
|---|---|
| **Open Discussion after AM Break** | **Various Identified Below** |

**Paul Watkins MD (UNC)**
Ximelagatran (XG) was pulled from the market before US approval. If you had a test to predict who would get hepatotoxicity, how predictive would it have to be to put it into practice?

**Tommy Andersson PhD (AstraZeneca)**
We didn't get that far.

**Neil Kaplowitz MD (USC)**

XG is unique in that even after short-term exposure, there is some lag before the appearance of hepatotoxicity. You wonder about immune mechanisms in that setting. In early development you need to look several weeks after exposure to assess for hepatotoxicity.

**Unknown Person**
Did you do lymphocyte stimulation tests in patients who had hepatotoxicity?

**Tommy Andersson PhD (AstraZeneca)**
Lymphocyte stimulation tests (LST) were done in patients who experienced ALT increases and in matched controls. However, the sampling of blood was done long after treatment had stopped, which may have caused the weak responses. However, the few positive LST tests were found in the ALT group.

**Jerry Stern MD (Boehringer-Ingelheim)**
You can change the drug metabolism phenotype by changing the gut flora. There may not be a stable metabolism phenotype especially with CYP450 metabolism.

**Clarification with response by Tommy Andersson PhD (AstraZeneca)**
There was a clarification about how many XG patients were rechallenged. Dr. Andersson replied that rechallenge wasn't intended, but that 18 patients were "reintroduced". Among those 18 patients, there was 1 positive rechallenge.

**Gerry Kenna PhD (AstraZeneca)**
Is it cost effective to do these studies (mechanistic studies, I think) for rare events? There is value in understanding mechanisms and susceptibilities.

**Tommy Andersson PhD (AstraZeneca)**
No animal models available (no animal showed hepatotoxicity with XG), so it is difficult to do studies.

**Neil Kaplowitz MD (USC)**
Would really severe cases of hepatotoxicity with XG have been identified by the genetic test?

**Tommy Andersson PhD (AstraZeneca)**
I don't know.

**Larry Lesko PhD (FDA/CDER)**
With the consortium (on serious AEs) you can test the performance (of a genetic test). Can you improve sensitivity of a test by selecting a subgroup for testing (e.g., a subgroup of patients who show early signs of toxicity)? Combining genetic test with ALT or TB might increase sensitivity, thereby increasing practicality and cost-effectiveness

**Frank Sistare PhD (Merck)**
The late check on liver enzymes- is that requirement in the guidance driven by the XG experience?

**John Senior MD (FDA/CDER)**
It is difficult to make a firm rule about late follow-up on liver enzymes

**Neil Kaplowitz MD (USC)**
One needs to do the late follow-up on liver enzymes early in development.

**Unknown Person**
Check after 5 half-lives.

**Will Lee MD (UTSW)**
Regarding ximelagatran (Exanta®), there were 8% of patients in the pre-approval trials who had 3X ULN transaminase elevations, but few serious outcomes. One patient died of fulminant HBV, 2 died of GI bleeds while on steroids for treatment of hepatotoxicity (their liver function was recovering). Ximelagatran is "cleaner" to use than coumadin. The advisory committee wanted to know what the real incidence of severe DILI with ximelagatran was but it was difficult to say from the limited data, even after 5,600 exposures. The drug was not approved by FDA but had been approved for short term use in the EU. It was subsequently withdrawn by the company because of concerns related to hepatotoxicity occurring as long as several weeks after the drug had been discontinued. There were still virtues to this drug since Coumadin has very significant bleeding complications associated with it, and many patients cannot take Coumadin at all because of inability to manage the dose adjustments necessary.

**Bob Temple MD (FDA/CDER)**
Do a head-to-head study of XG vs. coumadin and follow for bad outcomes (hepatotoxicity with XG vs. serious bleeding with coumadin).

| **Combining research with clinical study without regulatory penalty** | **Janet Woodcock MD (FDA)** |
|---|---|

No summary recorded

| **Open Discussion after Dr. Woodcock's talk** | **Various Identified Below** |
|---|---|

**John Senior MD (FDA/CDER)**
Your proposal will need buy-in from both Industry and FDA review divisions.

**Janet Woodcock MD (FDA)**
Everyone needs to participate in the plan to achieve a common goal. One study won't give a valid biomarker; need to repeat again and again. A sub-study could be done in patients with baseline liver abnormalities. Industry will buy-in if it is worth it to them.

**Bob Temple MD (FDA/CDER)**
One concern is that a test (new biomarker) will be used prematurely based on a subgroup analysis to identify the group with the problem, saying "if you use this test, you'll be okay".

**Janet Woodcock MD (FDA)**
We will be lucky to get this far.

**Unknown Person**
Develop a standard panel of biological samples that are taken when a patient reaches a certain threshold (>3x ULN). If you put this practice into the protocol, a company could build a repository.

**Frank Sistare PhD (Merck)**

Excellent guidance on toxicogenomics. Focus on preclinical aspect. Industry is still in the position in non-GLP studies to pick the best compound. However, there is hesitation to bring putative biomarkers into GLP studies because the process of validating a biomarker is ambiguous. It would be helpful if FDA could say if a longer term rodent study was done then there is no longer a need to go back to prior study

1. criteria for ignoring old study if a new study is done
2. clarify process of how a biomarker goes from exploratory to probable to valid

**Kate Gelperin (FDA/CDER)**
Earlier today there were different opinions about whether a patient should be continued on an investigational drug if they went beyond 3x ULN of ALT. Would this option be more acceptable if an additional specific informed consent was obtained from subjects in clinical trials who develop early signs of liver injury but who are willing to continue taking study drug under close surveillance and intense testing?

**Janet Woodcock MD (FDA)**
It depends on the setting of the trial. If there were many other therapeutic options, the patient might want to opt for something else; however, they might want to stay in the trial if nothing else is available. Re-consenting subjects in clinical trials who are willing to undergo additional testing while continuing study drug after showing early signs of liver injury may be appropriate, and may constitute a sub-study in some cases.

**Unknown Person**
What are the regulatory requirements for a drug that studied a biomarker while it was early in validation once it becomes accepted as valid?

**Jack Uetrecht PhD (U Toronto)**
It seems backward to look for biomarkers in new drugs. Shouldn't we go back and look at old drugs? How are we going to get funding to look at these biomarkers and who is going to study them? There is no encouragement for basic research in this field.

**Janet Woodcock MD (FDA)**
You are posing two separate questions:

1. old drugs – limited funding; check with NIH foundation
2. during new drug development – looking for new biomarkers; evolution of signal during trial

With enough leadership, we should be able to crystallize essential research questions

**Bob Temple MD (FDA/CDER)**
Further question- who's at risk? Look at current drugs- why do some people do better and some worse?